

受領書

平成17年 3月16日

特許庁長官

識別番号 100086243

氏名(名称) 坂口 博 様

提出日 平成17年 3月16日

以下の書類を受領しました。

項番	書類名	整理番号	受付番号	出願番号通知(事件の表示)
1	国際出願	JP920040033	50500472986	PCT/JP2005/ 4653

以上

特許協力条約に基づく国際出願書類

紙面に準拠し (注意: 電子データが標準となります)

0	受理番号記入欄	
0-1	国際出願番号	
0-2	国際出願日	
0-3	(受付所)	
0-4	様式-PC/101/01 この特許協力条約に基づく国際出願書類は、 0-1-1 当該によって作成された。	JPO-PAS 0322
0-5	申立て 出願人は、この国際出願が特許協力条約に従って処理されることを希望する。	
0-6	出願人によって指定された受理官庁	日本国特許庁 (RO/JF)
0-7	出願人又は代理人の登録番号	JP920040033
1	発明の名称	コンピュータシステム、これを構成するサーバ、その ジョブ実行制御方法及びプログラム
11	出願人 11-1 この欄に記載した者は 11-2 右の限定欄についての出願人である。	出願人である (applicant only) 米国を除く全ての指定国 (all designated States except US)
11-4a	名称	インターナショナル・ビジネス・マシーンズ・コーポ レーション
11-4m	Number	INTERNATIONAL BUSINESS MACHINES CORPORATION
11-5a	かて名	10504 アメリカ合衆国 ニューヨーク州アーモンク ニュー オーチャード ロード
11-5b	Address	New Orchard Road, Armonk, New York 10504 United States of America
11-6	国名 (国名)	アメリカ合衆国 US
11-7	住所 (国名)	アメリカ合衆国 US
11-11	出願人登録番号	390009531

特許協力条約に基づく国際出願願書

紙面による取扱い(注意:電子データが原本となります)

IB-1	その他の出願人又は発明者	出願人及び発明者である (applicant and inventor) 米国のみ (US only) 米澤 隆 YONEZAWA takashi 2428502 日本国 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ビ ー・エム株式会社大和事業所内 c/o Yamato site, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 2428502 Japan 日本国 JP 日本国 JP
IB-1-1	この欄に記載した者は	
IB-1-2	右の指定期限についての出願人である。	
IB-1-1a	氏名(姓名)	
IB-1-1en	Name (LAST, First):	
IB-1-5a	あて名	
IB-1-5en	Address:	
IB-1-6	国名(国名)	
IB-1-7	住所(国名)	
IB-2	その他の出願人又は発明者	出願人及び発明者である (applicant and inventor) 米国のみ (US only) 杉本 和敏 SUGIMOTO kazutoshi 2428502 日本国 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ビ ー・エム株式会社大和事業所内 c/o Yamato site, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 2428502 Japan 日本国 JP 日本国 JP
IB-2-1	この欄に記載した者は	
IB-2-2	右の指定期限についての出願人である。	
IB-2-1a	氏名(姓名)	
IB-2-1en	Name (LAST, First):	
IB-2-5a	あて名	
IB-2-5en	Address:	
IB-2-6	国名(国名)	
IB-2-7	住所(国名)	

特許協力条約に基づく国際出願願書

(紙面による写し (注意: 電子データが原本となります))

III-3 III-3-1 III-3-2 III-3-3a III-3-3en III-3-3ja III-3-3en III-3-3ja	その他の出願人又は発明者 この欄に記載した者は 右の指名欄についての出願人である。 氏名(姓名) Name (LAST, First) あて名	出願人及び発明者である (applicant and inventor) 米国のみ (US only) 美園 和久 MISONO kazuhisa 2428502 日本国 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ビ ー・エム株式会社大和事業所内 c/o Yamato site, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 2428502 Japan 日本国 JP 日本国 JP
III-3-3en III-3-3ja III-3-6 III-3-7	Address: 別荘(別名) 住所(別名)	 日本国 JP 日本国 JP
IV-1 IV-1-1ja IV-1-1en IV-1-2ja IV-1-2en IV-1-3 IV-1-4 IV-1-6 IV-2 IV-2-1ja IV-2-1en V V-1 VI-1 VI-1-1 VI-1-2 VI-1-3	代理人又は共通の代表者、追加のあて名 下記の者は国際出願において右記のとおりに 出願人のために活動する。 氏名(姓名) Name (LAST, First) あて名 Address: 電話番号 ファクシミリ番号 代理人登録番号 その他の代理人 氏名 Name(s) 領の指定 この領番を付いてされた国際出願は、最期 に於いてある、国際出願の時点で採算さ れる全てのPCTは特約国を指定し、獲得しう るあらゆる特約国の保護を求め、及び該当する 場合には家と国内特許の両方を求める 国際出願となる。 先の国内出願に基づく優先権主張 出願日 出願番号 国名	代理人 (agent) 坂口 博 SAKAGUCHI Hiroshi 2428502 日本国 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ビ ー・エム株式会社大和事業所内 c/o Yamato site, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 2428502 Japan 046-215-3318 046-275-5296 100086243 筆頭代理人と同じあて名を有する代理人 (additional agent(s) with the same address as first named agent) 市位 嘉宏(100091568); 上野 剛史(100108501) ICHII Yoshihiro(100091568); UENO Takeshi(100108501) 2004年 03月 19日 (19 03. 2004) 2004-080396 日本国 JP

特許協会の集約に基づく国際出願願書

願書による事し(注意:電子データが輸入となります)

VI-2	発明権証明書送付の請求	VI-1		
上記の発明の書類のうち、発明の番号のものについては、出願書類の登録簿主を作成し(願書事務局へ)送付することと、発明書件に対して請求している。				
VII-1	特許された国際調査報告書(ISA)	日本国特許庁 (ISA/JP)		
VIII-1	準立て	準立て紙		
VIII-1	発明書の特許に附する準立て			
VIII-2	出願し及び特許を与えられる国際出願書における出願人の資格に附する準立て			
VIII-3	発明の優先権を主張する国際出願書における出願人の資格に附する準立て			
VIII-4	発明者である者の準立て(米国を指定国とする場合)			
VIII-5	不利にならない開示又は新規性喪失の例外に附する準立て			
IX	調査費	用紙の枚数	添付された電子データ	
IX-1	願書(準立てを含む)	5		✓
IX-2	明細書	12		✓
IX-3	請求の範囲	4		✓
IX-4	要約	1		✓
IX-5	図面	6		✓
IX-7	合計	28		
IX-10	市町村印	添付		添付された電子データ
IX-6	手数料計算用紙	-		✓
IX-11	特許委任状の事し	-		✓
IX-17	PCT-SAFEと電子出願	-		-
IX-19	特許法とともに添付する図の番号	3		
IX-20	国際出願の使用言語名	日本語		
X-1	出願人、代理人又は代表者の署名押印	/100086243/		
X-1-1	氏名(姓名)	坂口 博		
X-1-2	署名者の氏名			
X-1-3	捺印			
X-2	出願人、代理人又は代表者の署名押印	/100091568/		
X-2-1	氏名(姓名)	市位 嘉宏		
X-2-2	署名者の氏名			
X-2-3	捺印			
X-3	出願人、代理人又は代表者の署名押印	/100108501/		
X-3-1	氏名(姓名)	上野 剛史		
X-3-2	署名者の氏名			
X-3-3	捺印			

特許協力の条約に基づく国際出願履歴表

(紙面による享し(注意: 電子データが原本となります))

受理官庁記入欄

10-1	国際出願として提出された書類の受理の日	
10-2	公開	
10-2-1	受理された	
10-2-2	不正開示がある	
10-3	国際出願として提出された事項を補充する 情報又は追加であつてその公開前に提出 されたものの公開の意思の日(訂正日)	
10-4	特許協力の条約第11条第1項に基づく発明補 正の期間内の発明の日	
10-5	出願人により特定された国際調査機関	ISA/JP
10-6	請求手続料未払いにつき 国際調査機関 に調査用紙を送付していない	

国際事務局記入欄

11-1	記録原本の受理の日	
------	-----------	--

PCT手数料計算用紙(領書付風書)

紙面による写し(注意:電子データが原本となります)
 [この用紙は、国際出願の一部を構成せず、国際出願の用紙の一枚に移入しない]

0	受理官庁記入欄	
0-1	国際出願番号	
0-2	受理官庁の日付印	
0-1	形式-PCT, RG/101 (付属書) このPCT手数料計算用紙は、 右記によって作成された。	JP0-PAS 0322
0-9	出願人又は代理人の登録番号	JP920040033
2	出願人	インターナショナル・ビジネス・マシーンズ・コーポレーション
12	所定の手数料の計算	金額/係数
12-1	送付手数料	T
12-2	領費手数料	S
12-3	領事出願手数料 (最初の30枚まで)	123200
12-4	30枚を越える用紙の枚数	0
12-5	用紙1枚の手数料	0
12-6	合計の手数料	0
12-7	$S1 + S2 =$	123200
12-12	50% electronic filing fee reduction	R
12-13	国際出願手数料の合計 (T+S+R)	1
13-17	納付するべき手数料の合計 (T+S+R)	206800
12-19	支払方法	送付手数料: 予納口座引き落としの承認 調査手数料: 予納口座引き落としの承認 国際出願手数料: 銀行口座への振込み
12-20	予納口座 受理官庁	日本国特許庁 (RG/JP)
12-20-1	上記手数料合計額の請求に対する承認	✓
12-21	予納口座番号	024154
12-22	年月日	2005年 03月 16日 (16. 03. 2005)
12-23	署名欄	

明 細 書

コンピュータシステム、これを構成するサーバ、そのジョブ実行制御方法及びプログラム

技術分野

- [0001] 本発明は、グリッドコンピューティングに関し、特にコンピュータに対してジョブの割り当てを行うスケジューリングの方法およびこれを実現するシステム構成に関する。

背景技術

- [0002] 近年、ネットワークで接続されたヘテロジニアス(異機種混在)な情報システムを統合して利用する、グリッドコンピューティングと呼ばれるシステムが注目されている。このシステムでは、ネットワーク上の複数のコンピュータにおけるCPUパワーやデータストレージなどのコンピュータ資源が共有され、仮想的な1つの高性能コンピュータとして利用される。複数のコンピュータに並列処理を行わせることで、1台1台の性能は低くとも高速に大量の処理を実行することが可能となる。
- [0003] グリッドコンピューティングでは、システムに与えられたジョブ(プログラムの処理単位)の実行順序を、ジョブの特性や優先順位に応じて決定するスケジューリングが非常に重要である。異機種混在環境を前提としたグリッド環境では、性能の異なる複数のコンピュータに対してジョブを割り当てる必要がある他、割り当てるジョブの大きさも様々であり、また、ローカルユーザだけでなく、遠隔地にいるユーザからもジョブを与えられることから、様々なジョブの実行スケジュールを適切に設定する必要があるためである。
- [0004] グリッドコンピューティングにおけるジョブのスケジューリングの方式には、大きく分けてプッシュ(Push)タイプとプル(Pull)タイプとがある。
- プッシュタイプ(例えば、非特許文献1参照)のスケジューリング方式では、スケジューラにジョブの投入があった場合に、スケジューラが、そのジョブを処理すべきコンピュータを割り当て、そのコンピュータに対してジョブの実行をリクエストする。この場合にスケジューラは、各コンピュータの利用状況をモニタリングし、ジョブの負荷情報等に応じて最適なコンピュータとジョブの割り当てを行う。

[0005] このプッシュタイプのスケジューリング方式は、主にクラスタグリッドと呼ばれる、数百台程度のサーバ(コンピュータ)で構成され、それらを同一サイト内に配置するタイプのグリッドシステムに多く用いられる。またプッシュタイプのスケジューリング方式は、最適なコンピュータに対してSpecificにジョブを割り当てることができ、最適なスケジューリングが可能である。特に、各コンピュータの挙動にばらつきがあり、かつユーザがコンピュータを使用した場合にジョブがキャンセルされる環境では、そのような動作環境を加味した上で、最適なスケジューリングにより非常に効率の良い運用が期待できる。

[0006] 一方、プルタイプ(例えば、非特許文献2参照)のスケジューリング方式では、各コンピュータがジョブを実行可能な状態になったときにスケジューラに対してポーリング(Polling)という形でジョブのリクエストを行う。そして、このリクエストが行われた際に、実行すべきジョブがあれば、スケジューラはそのコンピュータに対してジョブを割り当て、コンピュータが割り当てられたジョブを実行する。実行すべきジョブが無い場合は、コンピュータは、一定時間の経過後に再度ポーリングを行う。

[0007] このプルタイプのスケジューリング方式は、非常に簡単な構成で実装することができるため、数千台規模のグリッドシステムで多く使われる。特に、インターネット上に構築されるグリッドの場合には、ネットワークの制限のために、このプルタイプのスケジューリング方式が用いられる。またプルタイプのスケジューリング方式は、コンピュータの管理情報が簡単で、多数のコンピュータに対して適用しやすい。そして、ポーリングによる自己最適化(余裕のあるコンピュータほど頻繁にポーリングを行う)により、ある程度の効率化は図れる。

[0008] 非特許文献1:Chris Smith, "Open Source Metascheduling for Virtual Organizations with the Community Scheduler Framework (CSF)", Technical Whitepaper, Platform Computing Inc. 2003年8月。

非特許文献2:Eric Korpela, Dan Werthimer, David Anderson, Jeff Cobb, Matt Lebofsky, "Massively Distributed Computing for SETI", Computing in Science & amp; Engineering, Vol: 3, Issue: 1, Jan.-Feb. 2001, Pages: 78-83.

発明の開示

発明が解決しようとする課題

[0009] 上述したように、グリッドコンピューティングにおける従来のスケジューリング方式は、ブッシュタイプの場合、最適なジョブの割り当てを行うことができる。

しかし、最適なスケジューリングを行うために各コンピュータの使用状況を管理することが必要である。このため、スケジューラの動作が複雑になり、システム構成の変更（ジョブを実行するコンピュータの増減等）に容易に対応することができない。

また、スケジューラから各コンピュータにアクセスしてジョブの実行をリクエストする形態であるため、ファイアウォールの内側にあるコンピュータなどに対しては適用できない。

[0010] 一方、ブルタイプならば、上述したように、スケジューラが必要とするコンピュータの管理情報が簡単であるため、システム構成の変更にも容易に対応でき、またコンピュータからのポーリングによってネットワーク上でのアクセスが行われるため、ファイアウォールの内部にあるコンピュータに対してもジョブを割り当てることが可能である。

しかし、ポーリングによる自己最適化はあるものの、ブルタイプのスケジューリング方式では、実行すべきジョブがあっても、コンピュータからのポーリングを待ってジョブの実行をリクエストするため、時間的なロスが発生してしまう。

また、ポーリングが行われた場合には、実行すべき複数のジョブのうち、そのポーリングを行ったコンピュータにとって最適なジョブを割り当てて、そのジョブの種類によっては、より相応しい他のコンピュータがある可能性もあり、必ずしもシステム全体から見て最も効率の高い最適なスケジューリングができるわけではない。

[0011] そこで本発明は、グリッドコンピューティングにおいて、コンピュータの種類や使用状況およびジョブの種類の両面から最適なスケジューリングを実現すると共に、システム構成の変更に対して容易に対応が可能なシステムおよびそのジョブ実行制御方法を提供することを目的とする。

また本発明は、ファイアウォール等によるアクセス制限のあるコンピュータを含むグリッドにおいても、ブッシュタイプのスケジューリング方式と同等の最適なジョブの割り当てを可能とすることを他の目的とする。

また本発明は、ブッシュタイプとブルタイプとを組み合わせたスケジューリング方式、

およびこれを用いたシステムを提供することをさらに他の目的とする。

課題を解決するための手段

[0012] 上記の目的を達成するため、本発明は、ネットワークを介して接続された複数のコンピュータによりグリッドコンピューティングを実現する、次のように構成されたコンピュータシステムとして実現される。すなわち、このコンピュータシステムは、ネットワーク上のコンピュータに対してジョブの実行要求を行うコンピュータであるセンターサーバと、センターサーバからの要求に応じてジョブを実行するコンピュータであるプロセスサーバとを備えて構成される。そして、このセンターサーバは、プロセスサーバに対して実行すべきジョブを割り当て、ジョブ実行リクエストを発行するスケジューラ部と、プロセスサーバの情報を管理すると共に、スケジューラ部により発行されたリクエストを受け付けて、このリクエストにかかるジョブを割り当てられたプロセスサーバの状況に応じて、このリクエストをプロセスサーバに送信するエージェント部とを備える。

[0013] ここで、より詳細には、エージェント部は、複数のプロセスサーバに対して1対1で対応させて設けられる。

また好ましくは、エージェント部は、対応するプロセスサーバの能力および動作状況に関する情報をプロセスサーバから取得して管理する。そして、スケジューラ部は、エージェント部に管理されている情報に基づいて、プロセスサーバに対するジョブの割り当てを行う。

[0014] また、エージェント部は、プロセスサーバとのアクセスタイプに基づき、プロセスサーバからのポーリングによるアクセスに応じてスケジューラ部から受け付けたリクエストを送信したり、自身が管理するタイミングで前記スケジューラ部から受け付けたリクエストを送信したりする。

より具体的には、ファイアウォールを介してセンターサーバに接続されるプロセスサーバが存在する場合に、センターサーバからファイアウォールを越えてプロセスサーバにアクセスすることができないので、プロセスサーバからのポーリングによるアクセスを待ってリクエストを送信する。

[0015] また、上記の目的を達成する他の本発明は、コンピュータによりグリッドコンピューティングシステムにおけるジョブのスケジューリングおよび当該ジョブの実行要求を行う

、次のようなジョブ実行制御方法としても実現される。すなわち、このジョブ実行制御方法は、記憶装置に格納されたシステムを構成しジョブを実行するプロセスサーバの能力に基づき、このプロセスサーバの動作状況に関わらずにジョブの割り当てを行うステップと、ジョブの割り当てられたプロセスサーバに対するジョブ実行リクエストを発行するステップと、発行されたジョブ実行リクエストを一時的に保持し、ジョブの割り当てられたプロセスサーバの動作状況に応じて、このプロセスサーバに送信するステップとを含むことを特徴とする。

- [0016] さらに本発明は、コンピュータを制御して上述したセンターサーバの機能を実現させるプログラム、あるいはコンピュータに上記のジョブ実行制御方法の各ステップに対応する処理を実行させるプログラムとしても実現される。このプログラムは、磁気ディスクや光ディスク、半導体メモリ、その他の記録媒体に格納して配布したり、ネットワークを介して配信したりすることにより提供される。

発明の効果

- [0017] 以上のように構成された本発明によれば、グリッドコンピューティングにおいて、スケジューラとプロセスサーバとの間の通信を中継するエージェント部をセンターサーバに設け、プロセスサーバのアクセスタイプの違いをエージェント部の制御によって吸収することにより、プロセスサーバのアクセスタイプの違いによらずに、コンピュータの種類や使用状況およびジョブの種類両面から最適なスケジューリングを行うことができる。そしてプロセスサーバごとにエージェント部を設けることにより、プロセスサーバを増減してシステム構成を変更することが容易にできる。
- [0018] また、本発明によれば、プロセスサーバのアクセスタイプの違いをエージェント部の制御によって吸収することにより、従来はプッシュタイプのスケジューリングを行っていたジョブの実行要求を直接受け付けられるプロセスサーバと、従来はプルタイプのスケジューリングを行っていたポーリングを行った後にジョブの実行要求を受け付けるプロセスサーバとをシステム内に混在させ、かつプッシュタイプのスケジューリング方式と同等の最適なジョブの割り当てを行うことができる。

図面の簡単な説明

- [0019] [図1]本実施形態によるグリッドコンピューティングシステムの全体構成を示す図であ

る。

〔図2〕本実施形態によるグリッドコンピューティングシステムにおけるセンターサーバおよびプロセスサーバを実現するのに好適なコンピュータ装置のハードウェア構成の例を模式的に示した図である。

〔図3〕本実施形態におけるセンターサーバの機能構成を示す図である。

〔図4〕本実施形態におけるボーリングを行うプロセスサーバの機能構成とPSエージェント部との関係を示す図である。

〔図5〕本実施形態におけるボーリングを行わないプロセスサーバの機能構成とPSエージェント部との関係を示す図である。

〔図6〕本実施形態における各PSエージェント部によるジョブの実行要求時の動作を説明するフローチャートである。

発明を実施するための最良の形態

〔0020〕 以下、添付図面を参照して、本発明を実施するための最良の形態（以下、実施形態）について詳細に説明する。

図1は、本実施形態によるグリッドコンピューティングシステムの全体構成を示す図である。

図1に示すように、本実施形態のシステムは、ジョブの割り当てを行うセンターサーバ（CS）100による割り当てにしたがって実際にジョブを実行するプロセスサーバ（PS）200とを備える。センターサーバ100とプロセスサーバ200とは、インターネットその他のコンピュータネットワークにて接続されている。このコンピュータネットワークは、通信プロトコルや、有線か、無線かといった通信形式を問わず、またファイアウォールやその他のアクセス制限を設けたものであっても良い。

〔0021〕 図2は、本実施形態によるグリッドコンピューティングシステムにおけるセンターサーバ100およびプロセスサーバ200を実現するのに好適なコンピュータ装置のハードウェア構成の例を模式的に示した図である。

図2に示すコンピュータ装置は、演算手段であるCPU（Central Processing Unit：中央処理装置）11と、M／B（マザーボード）チップセット12およびCPUバスを介してCPU11に接続されたメインメモリ13と、同じくM／Bチップセット12およびAGP（

Accelerated Graphics Port)を介してCPU11に接続されたビデオカード14と、PCI(Peripheral Component Interconnect)バスを介してM/Bチップセット12に接続された磁気ディスク装置(HDD)15、ネットワークインターフェイス16と、さらにこのPCIバスからブリッジ回路17およびISA(Industry Standard Architecture)バスなどの低速なバスを介してM/Bチップセット12に接続されたフレキシブルディスクドライブ18およびキーボード/マウス19とを備える。

- [0022] なお、図2は本実施形態を実現するコンピュータ装置のハードウェア構成を例示するに過ぎず、本実施形態を適用可能であれば、他の種々の構成を取ることができる。例えば、ビデオカード14を設ける代わりに、ビデオメモリのみを搭載し、CPU11にてイメージデータを処理する構成としても良いし、外部記憶装置として、ATA(AT Attachment)やSCSI(Small Computer System Interface)などのインターフェイスを介してCD-R(Compact Disc Recordable)やDVD-RAM(Digital Versatile Disc Random Access Memory)のドライブを設けても良い。

- [0023] 図3は、センターサーバ100の機能構成を示す図である。

センターサーバ100は、各プロセスサーバ200に対するジョブの割り当て(スケジューリング)を行うスケジューラ部110と、プロセスサーバ200の管理を行うと共にプロセスサーバ200に対するリクエストおよびレスポンスの送受信を中継するPSエージェント部120とを備える。PSエージェント部120は、グリッドコンピューティングシステムを構成するプロセスサーバ200ごとに設けられている。そして、スケジューラ部110は、このPSエージェント部120を介して各プロセスサーバ200にアクセスする。

- [0024] スケジューラ部110は、例えば図2に示したプログラム制御されたCPU11とメインメモリ13や磁気ディスク装置15等の記憶手段とで実現され、その具体的な機能として図3に示すように、PS能力検査部111と、最適PS選択部112と、ジョブ依頼部113とを備える。

PS能力検査部111は、プロセスサーバ200の能力を各プロセスサーバ200に対応するPSエージェント部120に問合せ、情報取得する。

最適PS選択部112は、PS能力検査部111により取得されたプロセスサーバ200

の能力の情報に基づき、ジョブに応じて最適なプロセスサーバ200を選択し、割り当てを行う。このジョブの割り当てにおける最適化のロジックは任意で良い。

ジョブ依頼部113は、最適PS選択部112において選択されたプロセスサーバ200のPSエージェント部120に対してジョブ実行を要求するリクエストを発行する。

- [0025] 本実施形態では、PSエージェント部120がスケジューラ部110とプロセスサーバ200との間の通信を中継し、スケジューラ部110によるジョブの実行要求をプロセスサーバ200に代わって受け付ける。そのため、PS能力検査部111の問い合わせ先やジョブ依頼部113のリクエストの発行先がPSエージェント部120となっているが、上述したように、スケジューラ部110の機能自体は、既存のブッシュタイプのスケジューラと変わらない。したがって、スケジューラ部110には、既存のグリッドコンピューティングシステムで利用されているスケジューラを用いることができる。

- [0026] PSエージェント部120は、例えば図2に示したプログラム制御されたCPU11とメインメモリ13や磁気ディスク装置15等の記憶手段とで実現され、その具体的な機能として図3に示すように、PS状況管理部121と、PS能力管理部122と、ジョブ受け付け部123と、ジョブ依頼部124と、ポーリング待ち受け部125とを備える。

PS状況管理部121は、対応するプロセスサーバ200とアクセスして、プロセスサーバ200の現在の動作状況を把握する。

PS能力管理部122は、プロセスサーバ200のジョブ実行能力に関する統計情報等を管理し、スケジューラ部110のPS能力検査部111からの問い合わせに応じて、管理している情報を返す。ここで、ジョブ実行能力に関する統計情報等とは、単にCPU自身の処理能力や記憶装置の記憶容量の静的な情報だけでなく、CPUに対する負荷の時間変動や動作傾向といった動的な内容を統計的に処理して得られた情報を含む。

PS状況管理部121およびPS能力管理部122に管理される情報は、PSエージェント部120が対応しているプロセスサーバ200から取得され、図2のメインメモリ13や磁気ディスク装置15等の記憶手段に格納される。

ジョブ受け付け部123は、スケジューラ部110のジョブ依頼部113から発行されたジョブ実行リクエストを受け付ける。

ジョブ依頼部124は、ジョブ受け付け部123により受け付けられたジョブ実行リクエストを対応するプロセスサーバ200に送信する。

ポーリング待ち受け部125は、プロセスサーバ200からのポーリングによる、ジョブの実行が可能であるという通知を受付ける。

- [0027] 詳しくは後述するが、本実施形態では、ブッシュタイプのスケジューラを用いながら、プロセスサーバ200からのポーリングによるアクセスを受け付けてジョブ実行リクエストを送信することができる。PSエージェント部120のポーリング待ち受け部125は、このポーリングの受け付けのために用いられるものであり、ポーリングを行わずにセンターサーバ100における所望のタイミングで送信されたジョブ実行リクエストを受け付けられるプロセスサーバ200のPSエージェント部120に対しては、必要な構成要素ではない。

- [0028] 次に、プロセスサーバ200の機能構成と、対応するPSエージェント部120との関係について説明する。

上述したように、本実施形態において想定されるプロセスサーバ200は、ポーリングを行わないものを行うものがある。

図4は、ポーリングを行うプロセスサーバ200の機能構成とPSエージェント部120との関係を示す図である。

プロセスサーバ200は、図2に示したようなコンピュータ装置をグリッドコンピューティングシステムにおけるプロセスサーバ200として機能させるためのPS組み込み部210を備えている。

- [0029] PS組み込み部210は、例えば図2に示したプログラム制御されたCPU11とメインメモリ13や磁気ディスク装置15等の記憶手段とで実現され、その具体的な機能として図4に示すように、PS状況監視部211と、PS状況通知部212と、ジョブ受け付け部213と、ジョブ実行部214と、ポーリング実行部215とを備える。

PS状況監視部211は、自装置（プロセスサーバ200）の現在の使用状況や資源の状況を監視し、情報を収集する。

PS状況通知部212は、PS状況監視部211において収集されたPSの使用状況や資源の状況に関する情報を、センターサーバ100のPSエージェント部120に通知す

る。PSエージェント部120では、PS状況管理部121およびPS能力管理部122がこの通知を受け付け、メインメモリ13や磁気ディスク装置15等の記憶装置に格納して管理する。PS状況通知部212からセンターサーバ100への通知は、定期的に行っても良いし、プロセスサーバ200の動作状況が変更された際に行うようにしても良い。また、センターサーバ100の対応するPSエージェント部120から任意のタイミングでプロセスサーバ200へ問い合わせても良い。

ジョブ受け付け部213は、センターサーバ100のPSエージェント部120のジョブ依頼部124から送信されたジョブ実行リクエストを受け付ける。

ジョブ実行部214は、ジョブ受け付け部213により受け付けられたジョブを、プロセスサーバ200の資源を用いて実行する。

ポーリング実行部215は、PS状況監視部211の監視等に基づき、プロセスサーバ200がジョブを実行可能な状況である場合に、その旨を示す通知をセンターサーバ100のPSエージェント部120に対して行う。PSエージェント部120では、ポーリング待ち受け部125がポーリング実行部215からの通知を受け付けて、ジョブ依頼部124にジョブ実行リクエストの送信を行わせる。

- [0030] なお、ポーリング実行部215は、プロセスサーバ200がポーリングを行わずにセンターサーバ100から送信されたジョブ実行リクエストを受け付けられる場合、すなわちファイアウォールの内側にある等のアクセス制限がない場合は、設ける必要はない。

図5は、ポーリングを行わないプロセスサーバ200の機能構成とPSエージェント部120との関係を示す図である。

- [0031] 次に、上記のように構成された本実施形態のグリッドコンピューティングシステムの動作について説明する。

上述したように、センターサーバ100のスケジューラ部110は、既存のブッシュタイプのスケジューラと同様であるので、実行すべきジョブが発生すると、センターサーバ100の管理下にある全てのプロセスサーバ200を対象として、当該ジョブの割り当てを行う。このとき、PSエージェント部120からプロセスサーバ200の能力や動作傾向等の統計情報等を取得し、これらの情報およびジョブの種類や特性に基づいて最適なスケジューリングを行う。そして、ジョブが割り当てられたプロセスサーバ200の動作

状況に関わらずにジョブ実行リクエストを発行し、そのプロセスサーバ200に対応するPSエージェント部120に送信してジョブの実行要求を行う。

- [0032] PSエージェント部120の動作は、プロセスサーバ200のアクセスタイプが、センターサーバ100からのジョブの実行要求を直接受け付けられるタイプか、ファイアウォールの内側等によりボーリングを行った後にジョブの実行要求を受け付けるタイプかによって、動作が異なる。

図6は、各PSエージェント部120によるジョブの実行要求時の動作を説明するフローチャートである。

図6に示すように、PSエージェント部120は、ジョブ受け付け部123によりスケジューラ部110からジョブ実行リクエストを受け付け(ステップ601)、当該PSエージェント部120に対応するプロセスサーバ200がジョブの実行要求を直接受け付けられるタイプならば、そのジョブ実行リクエストを直ちにプロセスサーバ200へ送信する(ステップ602、604)。

- [0033] 一方、当該PSエージェント部120に対応するプロセスサーバ200がボーリングを行った後にジョブの実行要求を受け付けるタイプならば、PSエージェント部120は、プロセスサーバ200によるボーリングが行われるまで待機し(ステップ602、603)、ボーリングが行われた後にスケジューラ部110から受け付けたジョブ実行リクエストをプロセスサーバ200へ送信する(ステップ604)。なお、PSエージェント部120は、プロセスサーバ200によるボーリングが行われた際に送信すべきジョブ実行リクエストを持っていない場合は、何ら処理を行わず、ジョブの受け付けおよび次のボーリングのタイミングを待ち受ける。

- [0034] プロセスサーバ200は、センターサーバ100からのジョブの実行要求を直接受け付けられるタイプである場合、センターサーバ100のPSエージェント部120からジョブ実行リクエストを受け付けると、そのリクエストに応じてジョブを実行し、実行結果をセンターサーバ100のPSエージェント部120に返す。

一方、プロセスサーバ200がボーリングを行った後にジョブの実行要求を受け付けるタイプである場合、そのプロセスサーバ200は、ジョブの実行が可能な動作状況となった時点でボーリングを行い、ジョブ実行リクエストが送られるのを待つ。そして、セ

ンターサーバ100における対応するPSエージェント部120から送信されたジョブ実行リクエストを受け付けたならば、そのリクエストに応じてジョブを実行し、実行結果をセンターサーバ100のPSエージェント部120に返す。ジョブ実行リクエストを受信しなかった場合は、一定時間経過後に再度ポーリングを繰り返す。

[0035] 以上のプロセスサーバ200の動作は、従来のグリッドコンピューティングシステムにおけるプロセスサーバの動作と同様である。ただし、上述したように本実施形態では、センターサーバ100において各プロセスサーバ200に対応するPSエージェント部120が、プロセスサーバ200のアクセスタイプに応じて、自身が管理するタイミングでジョブの実行要求を行うか、またはプロセスサーバ200からのポーリングによるアクセスを待ってジョブの実行要求を行うかを制御できる。すなわち、プロセスサーバ200のアクセスタイプの違いをPSエージェント部120の制御によって吸収できるため、本実施形態によるシステムは、センターサーバ100からのジョブの実行要求を直接受け付けられるプロセスサーバ200と、ファイアウォールの内側等にありポーリングを行った後にジョブの実行要求を受け付けるプロセスサーバ200とを混在させることができる。

[0036] また、本実施形態は、上記のようにポーリングを行った後にジョブの実行要求を受け付けるプロセスサーバ200を含む構成とすることができる。そして、この場合でも、スケジューラ部110は、PSエージェント部120を介してプロセスサーバ200に対するジョブの実行要求を行うので、プロセスサーバ200の動作状況の影響を受けずに、プロセスサーバ200の能力とジョブの種類や特性とに基づき最適なスケジューリングを行うことができる。

[0037] さらに本実施形態は、上述したように、プロセスサーバ200ごとに1対1で対応するPSエージェント部120をセンターサーバ100に設け、このPSエージェント部120に、対応するプロセスサーバ200に関する情報の管理およびリクエストやレスポンスの送受信の制御を行わせる。また、スケジューラ部110は、PSエージェント部120が管理するプロセスサーバ200の情報に基づいて、プロセスサーバ200に対するジョブの割り当てを行う。このため、プロセスサーバ200を増減してシステム構成を変更する場合にも、対応するPSエージェント部120を追加または削除することによって、容易に対応することができる。

請求の範囲

- [1] ネットワークを介して接続された複数のコンピュータによりクリッドコンピューティングを実現するコンピュータシステムにおいて、
ネットワーク上のコンピュータに対してジョブの実行要求を行うセンターサーバと、
前記センターサーバからの要求に応じてジョブを実行するプロセスサーバとを備え、
前記センターサーバは、
前記プロセスサーバに対して実行すべきジョブを割り当て、ジョブ実行リクエストを発行するスケジューラ部と、
前記プロセスサーバの情報を管理すると共に、前記スケジューラ部により発行された前記リクエストを受け付けて、当該リクエストにかかるジョブを割り当てられたプロセスサーバの状況に応じて当該リクエストを当該プロセスサーバに送信するエージェント部とを備えることを特徴とするコンピュータシステム。
- [2] 前記エージェント部は、複数の前記プロセスサーバに対して1対1で対応させて設けられたことを特徴とする請求項1に記載のコンピュータシステム。
- [3] 前記エージェント部は、対応する前記プロセスサーバの能力および動作状況に関する情報を当該プロセスサーバから取得して管理し、
前記スケジューラ部は、前記エージェント部に管理されている前記情報に基づいて、前記プロセスサーバに対する前記ジョブの割り当てを行うことを特徴とする請求項2に記載のコンピュータシステム。
- [4] 前記エージェント部は、少なくとも一部の前記プロセスサーバに対して、当該プロセスサーバからのポーリングによるアクセスに応じて前記スケジューラ部から受け付けた前記リクエストを送信し、少なくとも他の一部の前記プロセスサーバに対して、自身が管理するタイミングで前記スケジューラ部から受け付けた前記リクエストを送信することを特徴とする請求項1に記載のコンピュータシステム。
- [5] 少なくとも一部の前記プロセスサーバは、ファイアウォールを介して前記センターサーバに接続されており、

前記エージェント部は、前記ファイアウォールを介して接続されている前記プロセスサーバに対して、当該プロセスサーバからのポーリングによるアクセスに応じて前記スケジューラ部から受け付けた前記リクエストを送信することを特徴とする請求項1に記載のコンピュータシステム。

- [6] グリッドコンピューティングシステムにおけるジョブのスケジューリングおよび当該ジョブの実行要求を行うサーバにおいて、
 システムを構成するコンピュータに対して実行すべきジョブを割り当て、当該ジョブの実行要求を行うスケジューラ部と、
 前記コンピュータの情報を管理すると共に、前記スケジューラ部による前記ジョブの実行要求を当該ジョブの割り当てられた当該コンピュータに代わって受け付け、当該コンピュータの状況に応じて当該ジョブの実行要求を当該コンピュータに対して行うエージェント部と
 を備えることを特徴とするサーバ。
- [7] 前記エージェント部は、システムを構成する前記コンピュータごとに設けられ、各々が対応するコンピュータとの間で設定された個別の通信形式で前記ジョブの実行要求を行うことを特徴とする請求項6に記載のサーバ。
- [8] 少なくとも一部の前記エージェント部は、システムを構成する前記コンピュータからのポーリングによるアクセスに応じて当該コンピュータに対する前記ジョブの実行要求を行い、少なくとも他の一部の前記エージェント部は、自身が管理するタイミングで対応する前記コンピュータに対する前記ジョブの実行要求を行うことを特徴とする請求項7に記載のサーバ。
- [9] グリッドコンピューティングシステムにおけるジョブのスケジューリングおよび当該ジョブの実行要求を行うサーバにおいて、
 システムを構成するコンピュータの能力および動作状況に関する情報を管理すると共に、当該コンピュータとの間で行われる通信を中継して、当該コンピュータの動作状況に応じて送受信を行うエージェント部と、
 前記エージェント部に管理されている前記情報に基づいて前記コンピュータに実行させるジョブの割り当てを行い、ジョブを割り当てられたコンピュータに対する当該ジョ

ブの実行要求を、当該エージェント部を介して行うスケジューラ部とを備えることを特徴とするサーバ。

- [10] 前記エージェント部は、システムを構成する前記コンピュータごとに設けられ、
 前記スケジューラ部は、ジョブを割り当てられた前記コンピュータに対応する前記エ
 ージェント部を介して当該ジョブの実行要求を行うことを特徴とする請求項9に記載の
 サーバ。
- [11] 前記スケジューラ部は、前記エージェント部に格納されている前記コンピュータの能
 力に関する情報に基づいて前記ジョブの割り当てを行い、かつ当該ジョブの割り当て
 られたコンピュータの動作状況に関わらず、当該ジョブの実行要求を行い、
 前記エージェント部は、少なくとも一部の前記コンピュータに対して、当該コンピ
 ュータからのポーリングによるアクセスに応じて前記スケジューラ部によるジョブの実行
 要求を送信し、少なくとも他の一部の前記コンピュータに対して、自身が管理するタイ
 ミングで前記スケジューラ部によるジョブの実行要求を送信することを特徴とする請求
 項9に記載のサーバ。
- [12] コンピュータにより、グリッドコンピューティングシステムにおけるジョブのスケジューリ
 ングおよび当該ジョブの実行要求を行うジョブ実行制御方法であって、
 前記コンピュータが、記憶装置に格納されたシステムを構成しジョブを実行するプロ
 セスサーバの能力に基づき、当該プロセスサーバの動作状況に関わらずにジョブの
 割り当てを行うステップと、
 前記コンピュータが、前記ジョブの割り当てられたプロセスサーバに対するジョブ実
 行リクエストを発行するステップと、
 前記コンピュータが、発行された前記ジョブ実行リクエストを一時的に保持し、前記
 ジョブの割り当てられた前記プロセスサーバの動作状況に応じて、当該プロセスサー
 バに送信するステップと
 を含むことを特徴とするジョブ実行制御方法。
- [13] コンピュータに、
 グリッドコンピューティングシステムを構成しジョブを実行するプロセスサーバの情報
 を記憶手段に格納して管理する機能と、

前記プロセスサーバの情報に基づき、当該プロセスサーバに対して実行すべきジョブを割り当て、ジョブ実行リクエストを発行する機能と、

発行された前記リクエストを受け付けて、当該リクエストにかかるジョブを割り当てられたプロセスサーバの動作状況に応じて当該リクエストを当該プロセスサーバに送信する機能と

を実現させることを特徴とするプログラム。

[14] 前記ジョブ実行リクエストを発行する機能では、前記コンピュータに、前記プロセスサーバの動作状況に関わらず前記ジョブの割り当てを実行させることを特徴とする請求項13に記載のプログラム。

[15] 前記リクエストを前記プロセスサーバに送信する機能では、前記コンピュータに、少なくとも一部の前記プロセスサーバに対して、当該プロセスサーバからのポーリングによるアクセスに応じて前記リクエストを送信させ、少なくとも他の一部の前記プロセスサーバに対して、自身が管理するタイミングで前記リクエストを送信させることを特徴とする請求項13に記載のプログラム。

[16] 前記リクエストを前記プロセスサーバに送信する機能では、前記コンピュータに、ファイアウォールを介して接続されている前記プロセスサーバに対して、当該プロセスサーバからのポーリングによるアクセスに応じて前記スケジューラ部から受け付けた前記リクエストを送信させることを特徴とする請求項13に記載のプログラム。

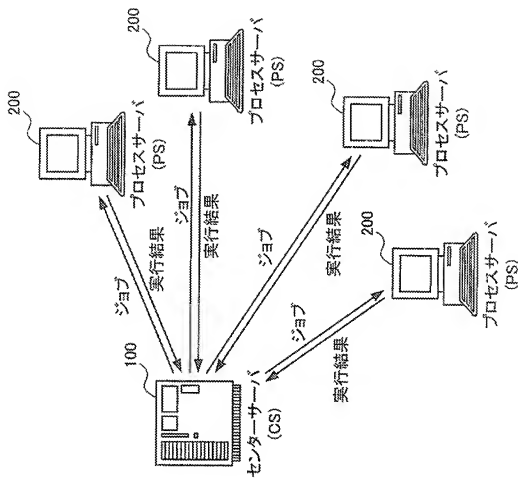
要 約 書

【要約】

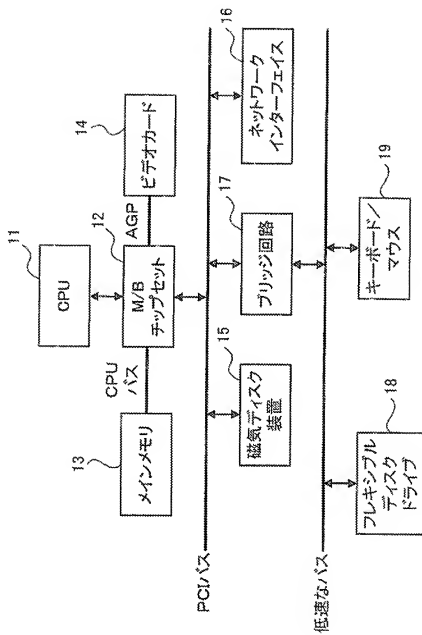
【課題】 グリッドコンピューティングにおいて、コンピュータの種類や使用状況およびジョブの種類の間から最適なスケジューリングが可能であり、システム構成の変更に対して容易に対応可能なシステムを実現する。〈BR〉

【解決手段】 グリッドコンピューティングにおいてネットワーク上のプロセスサーバに対してジョブの実行要求を行うセンターサーバ100が、プロセスサーバに対して実行すべきジョブを割り当て、ジョブ実行リクエストを発行するスケジューラ部110と、プロセスサーバ200の情報を管理すると共に、スケジューラ部110により発行されたリクエストを受け付けて、このリクエストにかかるジョブを割り当てられたプロセスサーバ200の動作状況やアクセスタイプに応じて、このリクエストをプロセスサーバ200に送信するPSエージェント部120とを備える。

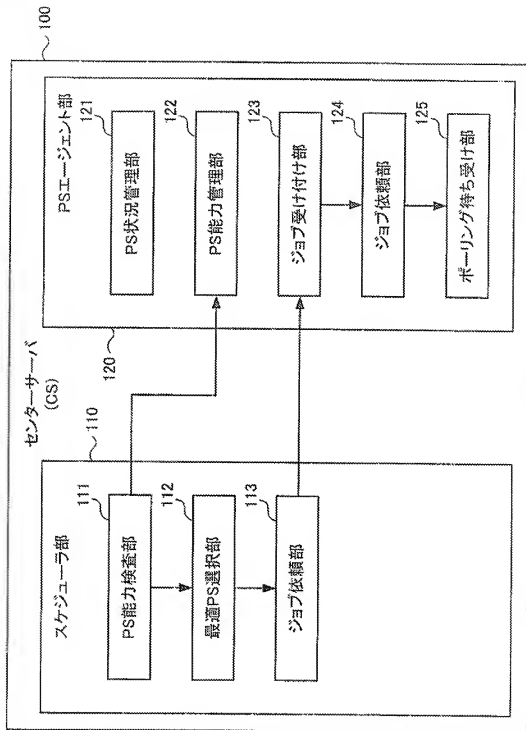
図1



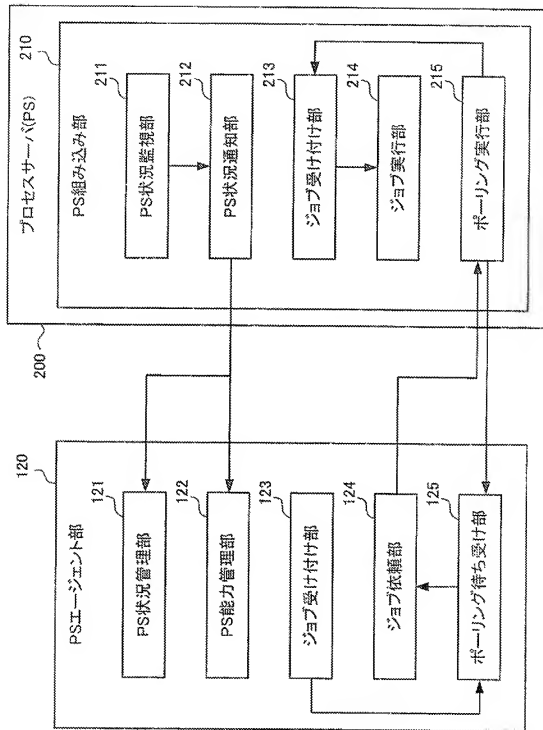
[図2]



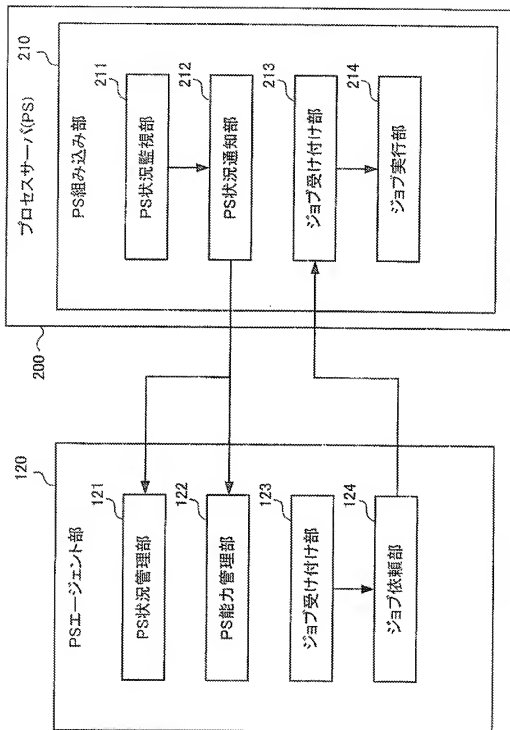
[図3]



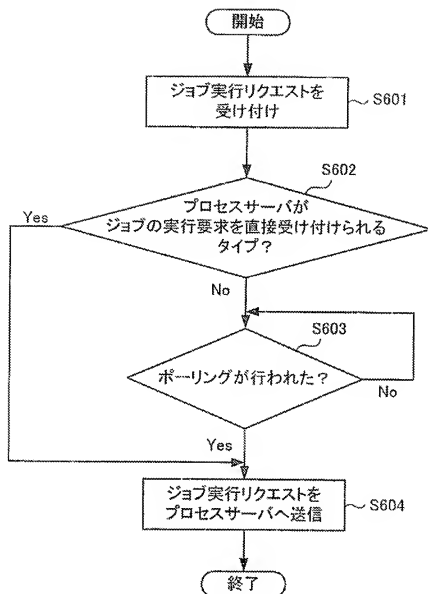
[図4]



[図5]



[図6]



包括委任状 (訳文)

アメリカ合衆国 1 0 5 0 4、ニューヨーク州アーモンク ニュー
オーチャード ロード のインターナショナル・ビジネス・マシー
ンズ・コーポレーション (以下、I B M と称する) は、日本国神奈
川県の弁理士 坂口博、弁理士 市位 嘉宏、弁理士 上野 剛史
を、I B M の適法な弁理士を定め、該弁理士に以下の事項をなす委
任された権限を与える。

1. I B M のために、特許協力条約に基づく国際出願に関す
る全ての事項について代理すること。
2. 上記国際出願の国内出願または指定国を取り下げるこ
と。
3. 上記国際出願に関連する国際予備審査請求に関する全て
の事項について代理し、上記指定国の要求または選択を
取り下げること。

2 0 0 2 年 7 月 9 日

インターナショナル・ビジネス・マシーンズ・
コーポレーション

ジェフリー・エル・フォアマン

委任状(PCT)

供題による写し (注意: 電子データが原本となります)

GENERAL POWER OF ATTORNEY

INTERNATIONAL BUSINESS MACHINES CORPORATION (hereinafter referred to as IBM) of New Orchard Road, Armonk, New York 10504, United States of America, does hereby appoint, HIROSHI SAKAGUCHI, YOSHIHIRO ICHII and TAKESHI UENO, Patent Attorneys of Kanagawa, Japan, as IBM's lawful attorneys, and hereby grant the above named attorneys the power:

1. to act for IBM in all matters relating to international applications under the Patent Cooperation Treaty,
2. to withdraw the above identified international applications or the designated State, and;
3. to act in all matters relating to a demand for the international preliminary examination with respect to the above identified international applications and to withdraw the demand or the election of any designated State.

Dated this 9th day of July 2002

INTERNATIONAL BUSINESS MACHINES CORPORATION

By:


Jeffrey L. Forman